DOCUMENT RESUME

ED 336 422 TM 017 230

AUTHOR Rodgers, Natalie; And Others
TITLE High Stakes Minimum Skills Tests: Is Their Use Increasing Achievement? ORE Publication Number 90.25.
INSTITUTION Austin Independent School District, Tex. Office of Research and Evaluation.
PUB DATE Apr 91
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Gains; *Achievement Tests; Analysis of Covariance; Grade 11; Graduation Requirements; High Schools; *High School Students; Longitudinal Studies; *Mathematics Achievement; *Minimum Competency Testing; Multivariate Analysis; School Districts; Testing Problems; Testing Programs; *Test Use
IDENTIFIERS Austin Independent School District TX; *Tests of Achievement and Proficiency; Texas Educational Assessment of Minimum Skills

ABSTRACT

Changes in mathematics achievement at grade 11 before and after the implementation of the Texas Educational Assessment of Minimum Skills (TEAMS) were studied in an urban school district in Texas. High stakes tests are used by state education agencies and school districts to monitor and increase educational achievement. The TEAMS is a minimum skills testing program; mastery of its exit level became a graduation requirement in 1987. Performance of students in the district was examined over a span of 5 years with 12,404 11th-graders (about 2,400 each year). For each subject, a percent correct score was determined for basic skills and high level skills categories of the Tests of Achievement and Proficiency. The grade point average of each student was included as a covariate to control for differential effects of instruction on students with varying previous achievement. Multivariate analysis of covariance was used. Over the period, basic skills scores did improve. It had been hypothesized that high level scores would decrease, as less attention was paid to higher achieving students when teachers concentrated on minimum performance. High level scores did not decrease, but they did not improve significantly. In addition, in the final year considered, after four years of concentration on the minimum competencies, both basic and high level scores declined. Implications for testing are discussed. One table and one graph present study data. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED336422

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NATALIE RODGERS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

High Stakes Minimum Skills Tests:

Is Their Use Increasing Achievement?

Natalie Rodgers
Vince Paredes
Evangelina Mangino, Ph.D.

Austin Independent School District
Office of Research and Evaluation
Austin, Texas

Paper presented at the Annual Meeting of the
American Educational Research Association
Chicago, April, 1991

ORE Publication Number 90.25

TM017230

BEST COPY AVAILABLE

High Stakes Minimum Skills Tests:
Is Their Use Increasing Achievement?

High stakes tests are used by state education agencies and school districts to monitor and increase overall achievement throughout the country. Tests which measure basic skills have been administered in Texas since 1980. The Texas Assessment of Basic Skills (TABS) was administered to students in grades 3, 5, and 9 from 1980-1985. The Texas Legislature, through House Bill 72, mandated a more comprehensive assessment plan. This new minimum skills testing program--the Texas Educational Assessment of Minimum Skills (TEAMS)--included new learning objectives and, according to the Texas Education Agency (TEA), "the rigor of the test was increased." The most important change in the TEAMS testing program was the addition of grade 11 as a graduation requirement.

This study analyzes changes in mathematics achievement at grade 11 before and after the implementation of the TEAMS at an urban school district in Texas. The first cycle of this test has ended and although TEA claims the objectives were met, TEA also admits that a dangerous narrowing of the curriculum has occurred to "tailor the curriculum" to the minimum skills measured by the test.

An unintended result of the excellence movement in education was an increase in the amount of state control of education. A prominent feature of the increase in state control has been the growth of state-mandated testing. (Anderson and Pipho, 1984) Most states have some form of minimum competency testing program. Programs differ among states in subject and grade level tested, the bearing of the test on graduation and promotion, and the kinds of preparations for and remediation available for nonmastery of the tests. This study also evaluates whether the mathematics scores for minimum skills and for high-level skills show the same trends for students with different overall achievement levels. This paper was written from the perspective of local public school evaluators who must report test results to school administrators and help them interpret the results.

The TEAMS tested students in mathematics, reading, and writing at grades 1, 3, 5, 7, and 9. At grade 11 students were tested in language arts and mathematics. Students who did not demonstrate mastery of the grade 11 test were allowed to retake the test until they mastered it. Students were offered as many as four opportunities during their junior and senior years to demonstrate mastery of the test. Beginning with students graduating in 1987, mastery of the exit-level test became a requirement for a diploma in Texas.

Educators attest that minimum competency testing programs focus the curriculum on minimum levels of achievement. Others say minimum competency testing has caused more changes in the teaching of basic skills to low-achieving students than the schools could have accomplished on their own. Minimum competency testing primarily affects low-achieving students. It tends to have little or no positive effect on those who are above average. It's only effect on above average achievers may be to suggest to them that the minimums are maximums, encouraging them to lower their expectations. (Anderson & Pipho, 1984)

There was fear and panic in school districts about the content of the new test required for graduation until the legislature, the State Board of Education and TEA issued objectives for each grade level tested. Objectives and measurement specifications on which the TEAMS was based were provided to teachers and other educators responsible for student instruction. The TEAMS Objectives and Measurement Specifications were very specific and explicit and included descriptions of skills, descriptions of test questions, descriptions of answer choices, and examples of test questions. Textbooks and curriculums were examined to see if they matched the skills, workshops were conducted at state, regional, and local levels on instructional activities for teaching the skills. Instructional administrators at the local level issued directives to teachers to concentrate on the skills measured by the TEAMS. Special area teachers were instructed to devote a certain amount of time each class to concentrating on the objectives. (Cruse, et al. 1985)

The TEAMS test "focused on only certain essential elements; in other words, it did not include many of the other concepts and skills" covered in the essential elements. Skills which required analytical thinking were not emphasized. Schools knew exactly which specific skills would be tested each year. Educators knew before each test exactly which of the essential elements would be tested. On the TEAMS, each objective was measured by the same number of questions. These objectives were more narrow than the Essential Elements which made up the curriculum in Texas.

TEA considered the TEAMS program as successfully meeting its purpose because the "overall student performance improved steadily in all areas." The success of the TEAMS program can be questioned because of the nature of the test and the way in which results are reported. The TEAMS has a very low ceiling. (Mangino, 1989) The skills included in the test, as the name indicates, are minimum skills and the results are reported in terms of the percentage of students who passed the test at a very low level (an approximate grade equivalent of 8.0 in mathematics, 8.4 in language arts). This grade equivalent was estimated through the equipercentile method equating TEAMS scores and TAP scores for students in the district.

As part of the rationale used by TEA for the changes made for the new testing cycle to begin in 1990-91, TEA has considered that "in the past few years concerns have arisen that in some districts the TEAMS objectives have become the primary focus of instruction and that the 'minimums' tested on the TEAMS have become the 'maximums.' The National Commission on Excellence in Education observed that minimum competency tests that are required for graduation cause the 'minimum to become the maximum and educational standards are lowered for all students. (Madaus, 1985) In some cases this practice has narrowed the curriculum to those objectives assessed on the TEAMS test. Teachers, instructional personnel may even be restricted in how skills tested by minimum competency tests are taught.

In 1986, a paper presented at AERA analyzing the effects of a state-mandated basic skills test (TABS) in Texas. Although this test had no repercussions upon whether students could graduate the test results had major repercussions because lawmakers, administrators, and the media used the results to compare and evaluate schools. The study warned of the dangers a high stakes test could have for high-achieving students and higher level skills in mathematics. (Mangino & Babcock, 1986)

Administrators and performance assessment evaluators used TEAMS scores as part of the information they used to evaluate teachers. TEA and the media used TEAMS scores to rank individual schools and districts. Theses rankings were published in newspapers across the state. (Ligon, 1989) Instructional time was diverted from the established curriculum to testing. Control of the curriculum shifted to the state with an overemphasis on basic skills. Accountability was based on TEAMS scores. (Ligon, 1989)

Districts designed programs to remediate all students, even students who had no problems mastering the TEAMS. Higher scores were achieved when students achieved mastery with four-out-of-four items measured, instead of three-out-of-four items. TEA assigned districts a grade through TEAMS. (Ligon, 1989 Johnstone) When a teacher's success is determined by how well students do on a test, the test determines the curriculum. If the test is narrow, measuring only minimum skills, the teacher may drill on those minimum skills, and the test becomes the "lowest common denominator" on which plans are based. The test determines the curriculum. (Ligon, 1989 Williams) The media used TEAMS scores to compare schools and Districts. Administrators used scores to compare schools. The latest TEAMS scores became headlines in local newspapers. Competition among schools within districts became fierce. Districts scrambled to find out the scores of comparable districts so they could respond to superintendents and communities who wanted to know why they ranked so poorly on TEAMS scores.

At the national level, the National Assessment of Educational Progress (NAEP) represents a minimum core of skills. The Office of Educational Research and Improvement (OERI) reported on NAEP's findings that "more students appeared to be gaining basic skills, but fewer were demonstrating a grasp of a the higher-level applications of those skills." Lower-level skills showed the most improvement. NAEP results did not show overall improvement in higher-order thinking skills and noted "some declines in the proportions of students who reached advanced levels." (OERI, 1990)

Content-validity asks the question, "Does the test measure what is taught? The question is reversed for minimum competency testing, "Does the curriculum measure what the test measures?" At the high school level, the basic skills test measures what a person needs to know to function in our society. How do these skills relate to the essential elements in Texas? (Mangino & Babcock, 1986)

Method

The Tests of Achievement and Proficiency (TAP) given at grade 11 were used to detect yearly changes in higher level mathematics skills. Students in grade 11 were selected because the minimum competency test required for graduation is first given and usually mastered by students at this grade. Items from the TAP were categorized into basic skills items and high level skills items. This was done by examining items on the TAP and determining whether they matched the objectives outlined for the TEAMS. Any items that did not fit TEAMS objectives were labeled high level items.

Performance of students in the district was examined over a span of five years. Subjects were 12,404 eleventh grade students (an average of about 2,400 per year) in the Austin Independent School District. Special education and limited English proficient students were not included in the study. Subjects took the TAP in April of each year included in the study.

The mathematics portion of the TAP test was used. It is believed that mathematics test scores are more subject to direct instruction and other factors than other areas. (Mangino, 1986) For example, reading scores are more subject to the effects of other courses, cultural factors and home conditions than are mathematics scores.

Design of the study

Two dependent variables were used in the analysis. For each subject a percent correct score was determined for both the basic skills and the high level skills categories.

The year tested was used as the independent variable. In addition, the grade point average of each student was included as a covariate to control for differential effects of instruction on students with varying previous achievement.

A multivariate analysis of variance (MANOVA) design was used. The effects used in the model were (1) the year of the test (2) GPA (3) year X GPA interaction and (4) GPA squared. The last effect was included because a previous study (Mangino and Babcock, 1986) indicated that the relationship between GPA and TAP scores might be curvilinear.

Hypothesis

It was expected that after the implementation of a new basic skills test (required for graduation) that (1) basic skills averages would show a significant increase and that (2) high level skills averages would show a significant decrease.

Analysis

The multivariate analysis was performed initially using both response variables (basic and high level scores). The MANOVA yielded a significant effect for year indicating that overall scores on the TAP varied between years. Since a significant main effect was found, the separate ANOVAs for basic and high level scores were examined (see attachment A).

The ANOVA using basic scores showed a significant main effect for year of test as expected, but the ANOVA using high level scores showed no main effect for year of test even with the large N involved in this study. This indicates that basic skills did change from year to year whereas high level skills did not change significantly from year to year.

The effect of the covariate was significant in both cases. The square of the covariate was also significant indicating an appreciable curvilinear component.

Results

A plot of the means by year of test (adjusted for the effect of other continuous variables in the model) reveals more information concerning the relationship between basic and high level skills. (See attachment B). Even though not statistically significant, high level skills scores show a rise and fall similar to basic skills scores over the five year period. In all cases when scores rise, basic skills gains are larger than high level score gains. And when scores in both categories drop, the high level scores show the larger drop.

The hypothesis that basic skills scores would increase was supported. The hypothesis that high level skills scores would decrease was not supported. However, high level scores did not increase significantly. Basic score changes were more positive than high level scores in years under study.

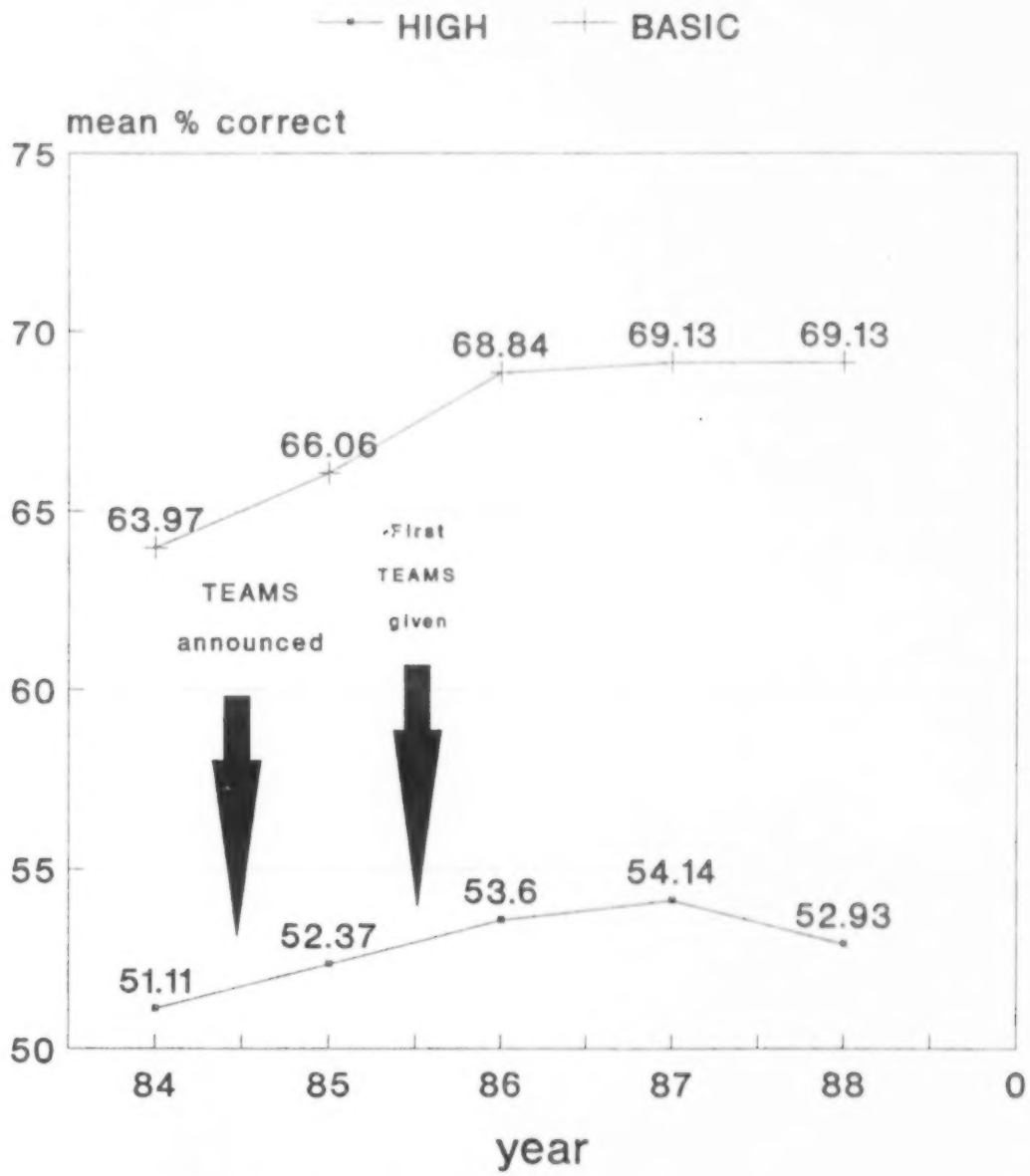
It is also interesting to note that after four years of emphasis on the minimum competency test, both basic and high level skills as measured by the TAP went down.

If a reduction of curriculum can be inferred from the decrease in high-level math scores, a serious warning must be issued to educational programs using test-driven curriculum programs such as the TEAMS. In districts or states where those programs are implemented, careful monitoring of higher-level skills must be done with other assessment instruments such as norm-referenced tests. Instructional coordinators must share the study's results with campus personnel to avoid limiting instruction to the skills being measured by the minimum skills testing program.

ANOVA results

Skill Level	Effect	p
HIGH	Year	.1823
	GPA	.0001
	Year x GPA	.1278
	GPA x GPA	.0001
BASIC	Year	.0454
	GPA	.0001
	Year x GPA	.2012
	GPA x GPA	.0001

Adjusted means



References

Anderson, B. & Pipho, C. (1984) State-mandated testing and the fate of local control. Phi Delta Kappan, 66 (3), 209-212.

Hays, William L. (1988). Statistics (4th ed.). New York: Holt, Rinehart and Winston, Inc.

Ligon, G., Johnstone, W., Mangino, E., Shaw, C., Brightman, M., Williams, P., Hoover, H.D., & Davis, E. (1989, January). Statewide testing in Texas. Symposium conducted at the annual meeting of the Southwest Educational Research Association, Houston, Texas. G. Ligon, (Chair) (Department of Management Information Monograph, Pub. No. 88.MO2). Austin Independent School District, Austin, Texas.

Madaus, G.E. (1985) Test Scores as Administrative Mechanisms in Educational Policy. Phi Delta Kappan, 66 (9), 611-617.

Mangino, E. (1986). Effects of the implementation of a minimum competency testing program on the acquisition of mathematics skills. (Doctoral dissertation, University of Texas, 1986) Dissertation Abstracts International, 47/06A, 2089.

Mangino, E. & Babcock, M.A. (1986, San Fransisco) Minimum competency testing: Helpful or harmful to high-level skills? Paper presented at the Annual Meeting of the American Educational Research Association.

Office of Educational Research and Improvement Bulletin, U.S. Department of Education Fall/Winter 1990

Popham, W.J., Cruse, K.L., Rankin, S.C., Sandifer, P.D., & Williams, P.L. (1985) Measurement-driven instruction: It's on the road. Phi Delta Kappan, 66 (9), 628-634.